Geospatial and Ecological Forces Shaping Multi-omics Shifts in Lycoris aurea (L'Hér.)

Sun Lei
leisun248@126.com
Resources of Hunan Province, College of Biological and Food Engineering, Huaihua University

Tan Zizhen
18670879067@163.com
Resources of Hunan Province, College of Biological and Food Engineering, Huaihua University

Tian Yuqiao
tyqhnhh@163.com
Hunan Provincial Engineering Research Center focuses on developing Polygonatum germplasm resources and advancing medicinal-food integration technologies.

Li Shenghua
lishenghua201@163.com
Resources of Hunan Province, College of Biological and Food Engineering, Huaihua University

Quan Miaohua
hhqmh100@163.com
Resources of Hunan Province, College of Biological and Food Engineering, Huaihua University

Research Article

Keywords:

DOI: https://doi.org/

Additional Declarations: No competing interests reported.

Title: Geospatial and Ecological Forces Shaping Multi-omics Shifts in *Lycoris aurea* (L'Hér.)

Sun Lei[1,2+], Tan Zizhen[1,2+], Tian Yuqiao[3], Li Shenghua[1,2*], Quan Miaohua[1,2*]

[1] Resources of Hunan Province, College of Biological and Food Engineering, Huaihua University, Huaihua 418000, China

[2] Hunan Provincial Higher Education Key Laboratory of Intensive Processing Research on Mountain Ecological Food, College of Biological and Food Engineering, Huaihua University, Huaihua 418008, China

[3] Hunan Provincial Engineering Research Center focuses on developing Polygonatum germplasm resources and advancing medicinal-food integration technologies. Huaihua 418008, China

[+] Co-first authors:

Sun Lei, Tan Zizhen

[*] Corresponding authors:

Li Shenghua: lishenghua201@163.com

Quan Miaohua: hhqmh100@163.com

Abstract

Background:

With stunning golden-yellow flowers, the perennial bulbous plant Golden Spider Lily [*Lycoris aurea* (L'Hér.)] thrives in temperate to subtropical regions of East Asia. Besides the ornamental value, *L. aurea* is also praised as a source of various biologics and a major pollinator attractant. The adaptability and resilience of *L. aurea* allow it to endure drought, waterlogging, and suboptimal soil conditions. However, the effect of habitat and ecology on *L. aurea* metabolites remains unexplored.

Method

This study investigates the spatial and environmental influences on the metabolites and gene expression of *L. aurea* using a combined metabolomic and transcriptomic approach. Bulbs were collected from five locations in China, selected for their diverse ecological and environmental conditions, including variations in temperature, precipitation, soil pH, and selenium concentration. The metabolites in the bulb extracts were analyzed using UPLC-MS/MS, and RNA sequencing was performed to capture gene expression data. The correlation between environmental factors and metabolite accumulation, as well as gene expression, was analyzed to understand the plant's adaptive mechanisms.

Results

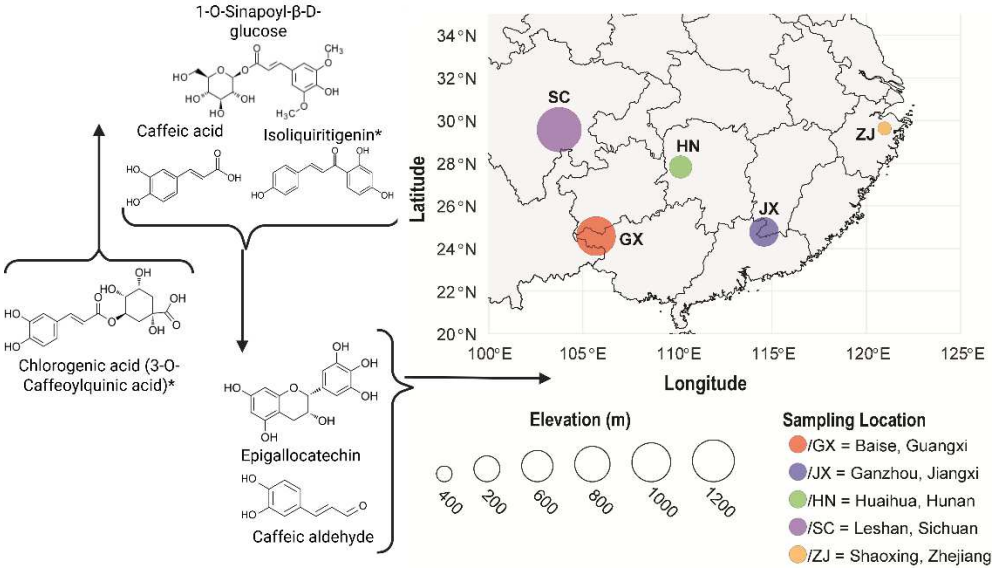
The analysis revealed significant variations in the metabolite and gene expression profiles of *L. aurea* bulbs from different locations. Alkaloids, flavonoids, and phenolic acids were among the most abundant metabolites identified, with geographic differences influencing their abundance. For example, higher-altitude samples exhibited increased levels of flavonoids and terpenoids, compounds associated with plant defense mechanisms. Gene expression patterns mirrored these findings, with upregulation of genes involved in secondary metabolite biosynthesis in plants from higher altitudes. Environmental factors such as soil pH, selenium concentration, and altitude were found to play significant roles in shaping both the metabolite composition and gene expression patterns of the plant. Metabolites related to plant defense, lignin production, biotic and abiotic stresses were affected mainly by spatial heterogeneity.

Conclusion

This study highlights the impact of environmental and spatial factors on the biochemical and genetic profiles of *L. aurea*, showing how the plant adapts to varying conditions. Altitude, latitude, and soil composition are key determinants of its metabolic output. These findings provide insights into optimizing cultivation and enhancing the medicinal properties of the plant. Future research should investigate the regulatory networks that link gene expression to metabolite biosynthesis in response to environmental cues.

Keywords: *Spatial, ecological, plant adaptation, golden spider lily, metabolomics, transcriptomics*

Graphical Abstract



1. Background

Lycoris aurea (L'Hér.), commonly known as the Golden Spider Lily, is a well-known member of the Amaryllidaceae family [1-4]. Native to East Asia, especially southern China, Taiwan, and Japan, this perennial bulbous plant is distinguished by its vivid golden-yellow blooms that blossom in late summer or early fall, often preceding its elongated, slender, leafless stems [3, 5]. The petals of its flower are arched backward like spider legs, which has led to its colloquial designation, 'Spider Lily' [5]. In addition to its visual allure, *L. aurea* has considerable cultural, medicinal, and ecological significance, making it a focal point in botanical and pharmacological research [3, 5]. In its habitat, the nectar-rich *L. aurea* flowers attract various pollinators, such as bees and butterflies, providing essential pollination services [6, 7]. In traditional Chinese medicine, *L. aurea* is highly valued for its rich composition of bioactive alkaloids [3]. For instance, the pulverized bulbs of *L. aurea* have been used as a poultice for treating burns, scalds, and ulcers in Hubei, China [8]. Additionally, *L. aurea* generates many bioactive alkaloids, notably Lycorine, Homolycorine, and other associated chemicals, with most research documenting them in its ovary [9] and bulbs [4, 5, 10-13]. These metabolites are responsible for the plant's biological activity and are of interest in traditional medicine, as they are believed to possess antitumor, antimicrobial, and anti-inflammatory properties. Lycorine showed antiviral and anticancer activity [14, 15], Galanthamine has anti-inflammatory properties [16] and alleviates the neuropsychiatric symptoms of Alzheimer's [17, 18]. Though other *Lycoris* sp. ubiquitously synthesizes these compounds, their amount varies between species [4] and location [3], which can be attributed to several genes from alkaloid and phenylpropanoid biosynthesis pathways such as Aldehyde Dehydrogenase (ALDH), Phenylalanine Ammonia-Lyase (PAL), and Norbelladine 4′-O-Methyltransferase (N4OMT) [9, 19, 20], as well as ecological conditions such as soil quality, sunlight, rainfall, and soil microbiota [4, 5, 7, 12, 21, 22].

L. aurea flourishes in temperate to subtropical climates, in shaded and humid habitats with optimal drainage and sufficient irrigation [4, 5]. They often grow in sheltered wet slopes near streams in mountainous regions, on the edges of woods, rice paddies, and plantations [22]. However, their distinctive pre-foliar blooming reduces nutrient requirements, enabling them to thrive in various other environments [3]. Additionally, its bulbous roots store nutrients and energy [5], allowing it to endure drought, waterlogging, and suboptimal soil conditions and to regenerate after each blooming season [22]. So, the herb is also well-suited to diverse altitudes, temperatures, light, water, and soil conditions, exhibiting resilience against pests and diseases [5]. Since soil composition affects the nutritional and metabolic profile of *L. aurea* bulbs [11], other agroecological, seasonal, and spatial factors may have a similar impact [3, 4, 23]. For example, suboptimal irrigation during the latter development stage of *L. aurea* was reported to enhance bulb alkaloid concentration [21]. Other factors affecting plant alkaloid production include age, microbial attack, and grazing [23]. For instance, with a decrease in latitude, herbivore and pathogen pressure on *Anguinaria canadensis* increases, while the diversity and toxicity of alkaloids increase [24]. In contrast, plants grown at higher altitudes, where temperatures are cooler and sunlight is more intense, may produce different concentrations or types of metabolites to adapt to harsher conditions [25]. Similarly, geographic differences in soil composition, humidity, and sunlight can influence the synthesis of specific alkaloids, leading to variations in the medicinal and toxic properties of the plant across regions [26].

Plants from niches with higher rainfall, such as the subtropical areas of Taiwan, may exhibit a different metabolic composition compared to those growing in drier, temperate zones in China or Japan [27]. When determining the best places to get bioactive molecules from, scientists must consider how different altitudes, latitudes, and longitudes affect the plant's metabolism [24, 28, 29]. Temperate, subtropical, and higher-altitude environments differ ecologically and climatically, leading to metabolic variations that affect the chemical composition of plants, as well as their growth patterns, blooming cycles, and overall health [23, 27, 30]. For instance, Lycorine, the principal alkaloid of *L. aurea,* is higher in temperate areas and increases with soil pH, moisture content, and Selenium (Se) levels, while those in humid or tropical climates may produce different alkaloids [3]. Studying spatial effects on a plant's metabolic differences will provide valuable insights into its adaptability and its secondary metabolites' role in defense [24]. Moreover, it could guide the development of optimal cultivation strategies to improve the yield of valuable bioactive chemicals [3]. By understanding how altitude, latitude, and soil composition affect a plant's metabolic output, we can tap into therapeutic potential while limiting toxicity [31]. Since *L. aurea* is used as a medicinal plant in Ayurvedic and Chinese medicine practice, the pharmacological properties of its various metabolites have been well studied, and some have even been licensed as contemporary therapeutics [21]. Conversely, despite a few restricted studies examining the effects of soil, water, altitude, and latitude on the plant, the influence of location on *L. aurea* remains mostly enigmatic [3, 12, 21, 31]. Therefore, to maintain the sustainability of the L. aurea ecosystem and cultivate it for medical purposes, we devised a method to understand how the plant adjusts its metabolism to different geographic situations by examining the metabolome and transcriptome of the plant collected from various locations in China and performing a conjoint analysis.

2. Method

2.1. Sample Collection

Sampling sites were selected from the natural habitats of *L. aurea* to capture a diverse array of variables affecting alkaloid production, encompassing spatial factors like latitude, longitude, and elevation [29, 31], alongside environmental parameters such as temperature, precipitation, soil pH, and Se concentration [25, 26, 32, 33]. We selected five sampling sites within the natural habitat of *L. aurea* in China, characterized by diverse variables (Fig. 1A). *L. aurea* bulbs were collected in triplicate from five locations in China between the 7th and 24th of August 2023 (Fig. 1B). Sampling locations included Baize city (105° 42′, 24° 35′, 877m) from Guangxi province (GX), Ganzhou city (114° 35′, 24° 47′, 446.5m) from Jiangxi province (JX), Huaihua city (110° 10′, 27° 49′, 277m) from Hunan province (HN), Lashan city (103° 44′, 29° 56′, 1168.9m) from Sichuan province (SC), and Shanxing city (120° 57′, 29° 38′, 189m) from Zhejiang province (ZJ) (Fig. 1C). Bulb samples were collected in accordance with the Regulations on Wild Plant Protection of the People's Republic of China, the Seed Law of the People's Republic of China, and relevant local regulations. Necessary permissions for collection were obtained from Hunan Provincial Forestry Bureau, with the permit number No. 222 of 2025. Specimens were identified by Prof. Zeng Hanyuan, and voucher specimens were deposited in the Huaihua University with accession number HHU2022-8. Collected ~~B~~bulbs were washed, flash-frozen, labelled, and stored at -80° C before extraction.

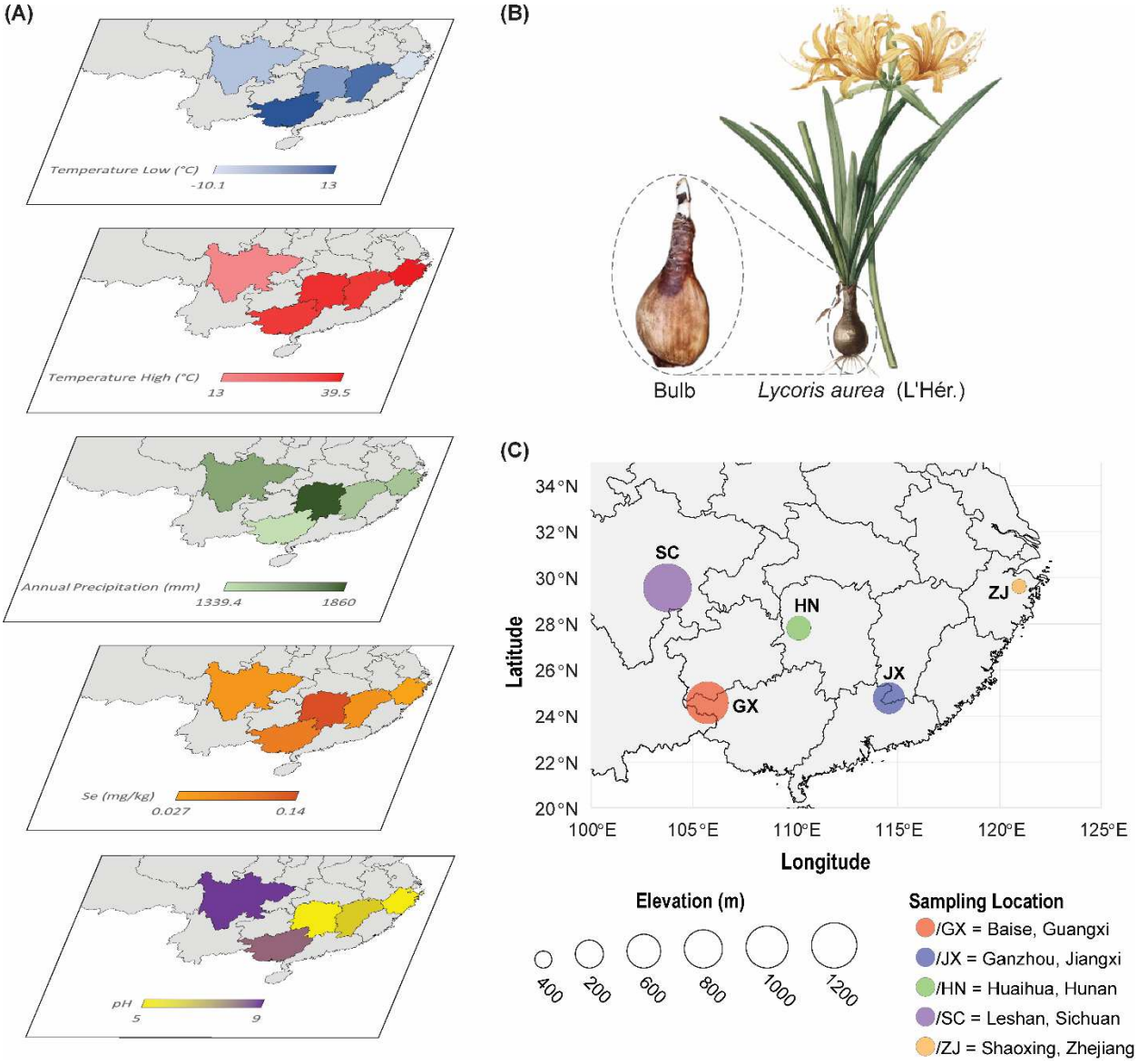


Fig. 1: Collection of *L. aurea* bulb samples.

(A) Spatial and environmental analysis of *L. aurea* natural niches in China revealed five suitable sampling locations characterized by diverse factors. Maps illustrate the lowest and highest temperatures (°C), annual precipitation (mm), selenium (Se) (mg/kg), and soil pH, which may influence alkaloid production in *L. aurea*. (B) Image of *L. aurea* (L'Hér.), or Golden Spider Lily, featuring its distinctive bulbous root. (C) The sampling sites of L. aurea are indicated by colored circles, with circle size reflecting elevation.

2.2.Metabolomics Analysis

2.2.1. Sample Preparation and Extraction

Bulbs were freeze-dried (Scientz-100F) and ground (30 Hz, 1.5 min) using a planetary ball mill (MM 400, Retsch). 50 mg of powder was weighed (MS105DM) and dissolved in 1200

μL of pre-cooled (−20 °C) 70% methanol aqueous internal standard extract. The mixture was vortexed every 30 min for 30 sec (6 times). After centrifugation (12000 rpm, 3 min), the supernatant was filtered through a 0.22 μm membrane and stored in an injection vial for UPLC-MS/MS analysis.

2.2.2. Metabolomic Analysis

Metabolites in *L. aurea* (L'Hér.) bulbs extracts were analyzed by Wuhan MetWare Biotechnology Co., Ltd. (Wuhan, China) using a UPLC-ESI-MS/MS system (UPLC, ExionLC™ AD, https://sciex.com.cn/) and a Tandem mass spectrometry (MS/MS) system (https://sciex.com.cn/), following their established protocols [34]. The UPLC system used an Agilent SB-C18 1.8 μm, 2.1 mm × 100 mm column with mobile phases A (ultrapure water with 0.1% formic acid) and B (acetonitrile with 0.1% formic acid). The gradient started at 5% B for 0 min, increased to 95% B by 9 min, and reverted to 5% B at 11.1 min. Flow rate: 0.35 mL/min; column temperature: 40°C; injection volume: 2 μL. ESI source temperature was set at 500°C, with ion spray voltages of +5500 V (positive) and -4500 V (negative). Gas I, II, and curtain were maintained at 50, 60, and 25 psi, respectively, under conditions of high collision-induced ionization. The QQQ scan in MRM mode utilized nitrogen as the collision gas, with optimized DP and CE for each ion pair.

2.2.3. Metabolomics Data Acquisition and Quantitative Analysis

Metabolite mass spectrometry data from various samples were analyzed, integrating the peak areas of all chromatographic peaks. The mass spectrometry peaks corresponding to the same metabolite across different samples were also integrated and corrected [35]. Mass spectrometry data were processed with Analyst 1.6.3 (AB SCIEX, Concord, Ontario, Canada). Metabolites from the samples were analyzed qualitatively and quantitatively via mass spectrometry, utilizing the local metabolic database. MultiQuant software (Framingham, Massachusetts, USA) is used to open the mass spectrometry file for the integration and correction of chromatographic peaks. Quality control (QC) analysis was conducted on samples prepared by mixing extracts, assessing repeatability through overlapping TIC graphs under identical processing methods.

Data were scaled using unit variance (UV) and analyzed via unsupervised principal component analysis (PCA) with the prcomp function in R (www.r-project.org). Hierarchical cluster analysis (HCA) of samples and metabolites was visualized as heatmaps with dendrograms, and Pearson correlation coefficients (PCC) between samples were calculated using the cor function in R and displayed as heatmaps. Both HCA and PCC analyses were performed using the R package ComplexHeatmap. In HCA, metabolite signal intensities, normalized and scaled to UV, were represented with a color spectrum.

Two groups were compared by identifying differential metabolites based on variable importance projection (VIP > 1) and absolute log fold change ($|Log_2FC| \geq 1.0$). VIP values were derived from orthogonal partial least squares discriminant analysis (OPLS-DA) using score and permutation plots, with the R package MetaboAnalystR. Data were log-transformed (log_2) and mean-centered before OPLS-DA. A permutation test (200 permutations) was performed to prevent overfitting.

2.2.4. Metabolite Annotation and Enrichment Analysis

Identified metabolites were annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) Compound database (http://www.kegg.jp/kegg/compound/) and mapped to

the KEGG Pathway database (http://www.kegg.jp/kegg/pathway.html). Pathways with significantly regulated metabolites were analyzed using metabolite set enrichment analysis (MSEA), and p-values from the hypergeometric test determined their significance.

2.3.Transcriptomics Analysis

2.3.1. RNA Extraction, cDNA Library Construction, and RNA-Sequencing

Three biological replicates from each location were used for RNA-sequencing analysis, performed by MetWare Biotechnology Co., Ltd. (Wuhan, China), following their standard protocol. Total RNA was extracted using the RNeasy Mini Kit (QIAGEN, Germany). RNA integrity was assessed by agarose gel electrophoresis, and RNA concentration was measured using the Qubit® RNA Assay Kit (Life Technologies, USA). RNA quality was verified using the Qsep400 Bioanalyzer (Bioptic, Taiwan). cDNA was synthesized from 1 μg of RNA per sample, and sequencing libraries were prepared using the NEBNext® UltraTM RNA Library Prep Kit for Illumina® (Nebraska, USA). The indexed samples were clustered on the cBot Cluster Generation System (Illumina®) and sequenced on an Illumina platform, generating 150 bp paired-end reads.

2.3.2. Transcriptome Data Acquisition and QC Analysis

The initial dataset was cleaned using the 'fastp' tools [36] to remove adapter sequences, reads with >10% N content, sequencing reads exceeding 10% of the base count, and reads with >50% low-quality bases (Q≤20). Clean reads were used for all subsequent analyses. Transcriptome assembly was performed using 'Trinity' (https://github.com/trinityrnaseq/trinityrnaseq) [37], and the 'Corset' tool [38] reorganized transcripts into 'UniGene' clusters. Potential coding regions (CDS) were identified using 'TransDecoder' (https://github.com/TransDecoder/).

2.3.3. Gene Functional Annotation

The Unigene sequence was compared with KEGG, Nr (NCBI non-redundant protein sequences), Swiss-Prot, GO, COG/KOG, and TrEMBL databases using DIAMOND [39] BLASTX. The amino acid sequence of the Unigene was predicted and further compared with the Pfam (Protein family) database using HMMER (http://hmmer.org/) to obtain annotation information.

2.3.4. Gene Expression Quantification and Differential Analysis

Gene expression levels were estimated using the RSEM software [40], and FPKM (Fragments Per Kilobase of transcript per Million fragments mapped) was calculated based on gene length. Differential expressions between biological replicates were analyzed using DESeq2 [41, 42], with the Benjamini-Hochberg method applied to correct for multiple hypothesis testing, yielding the false discovery rate (FDR). Differentially expressed genes (DEGs) were identified with $|\log_2\text{Fold Change}| \geq 1$ and $\text{FDR} < 0.05$.

Enrichment analysis was performed using the hypergeometric test, with KEGG and GO analysis based on pathways and terms, respectively. Transcription factor analysis was done using iTAK [43]. DEGs were subjected to DIAMOND BLASTX against the genome of a related species, and predicted protein-protein interactions (PPIs) were obtained from the STRING database (http://string-db.org/). Simple sequence repeats (SSRs) were identified using MISA [44], and primers were designed using Primer3 (http://primer3.sourceforge.net/releases.php). Weighted correlation network analysis (WGCNA) was performed using the WGCNA R Package [45].

2.4. Combined Metabolome and Transcriptome Analysis

Combined metabolome and transcriptome analyses were conducted on the DEGs and differentially accumulated metabolites (DAMs) to assess pathway enrichment. PCA was performed to visualize distinctions across sample groups, followed by KEGG pathway, KEGG enrichment, KGML, and Canonical Correlation Analysis (CCA) [46]. Expression trends and correlations were analyzed using gene-metabolite networks with a Pearson Correlation Coefficient (PCC) > 0.8 and p-value < 0.05 in each sample group [47].

3. Results

3.1. Spatial Influence on *L. aurea* Bulbs Metabolome

We analyzed the metabolomes of 70% methanolic extracts from L. aurea bulbs across 15 samples from 5 geographical locations in China, identifying 731 metabolites (Table S1). These included 36.11% Alkaloids, 9.3% Flavonoids, 5.75% Lignans and Coumarins, 18.19% Phenolic acids, 1.92% Quinones, 12.59% Terpenoids, and 16.14% other compounds (Fig. 2A). OPLS-DA analysis (Fig. S1) effectively discriminated groups based on measured features, with strong predictive ability (Q^2) and excellent model fit (R^2), identifying 546 differential accumulated metabolites (DAMs) with VIP > 1 and $|Log_2FC|$ ☒ 1.0 (Table S2). 3D PCA (Fig. 2B) showed significant variance among groups, with biological replicates clustering closely. The hierarchical clustered heat map (Fig. 2C) confirmed differential metabolite abundance and identified biologically relevant pathways. These findings suggest location-specific differences in the metabolomic profiles of *L. aurea* bulbs, potentially influenced by environmental or genetic factors.

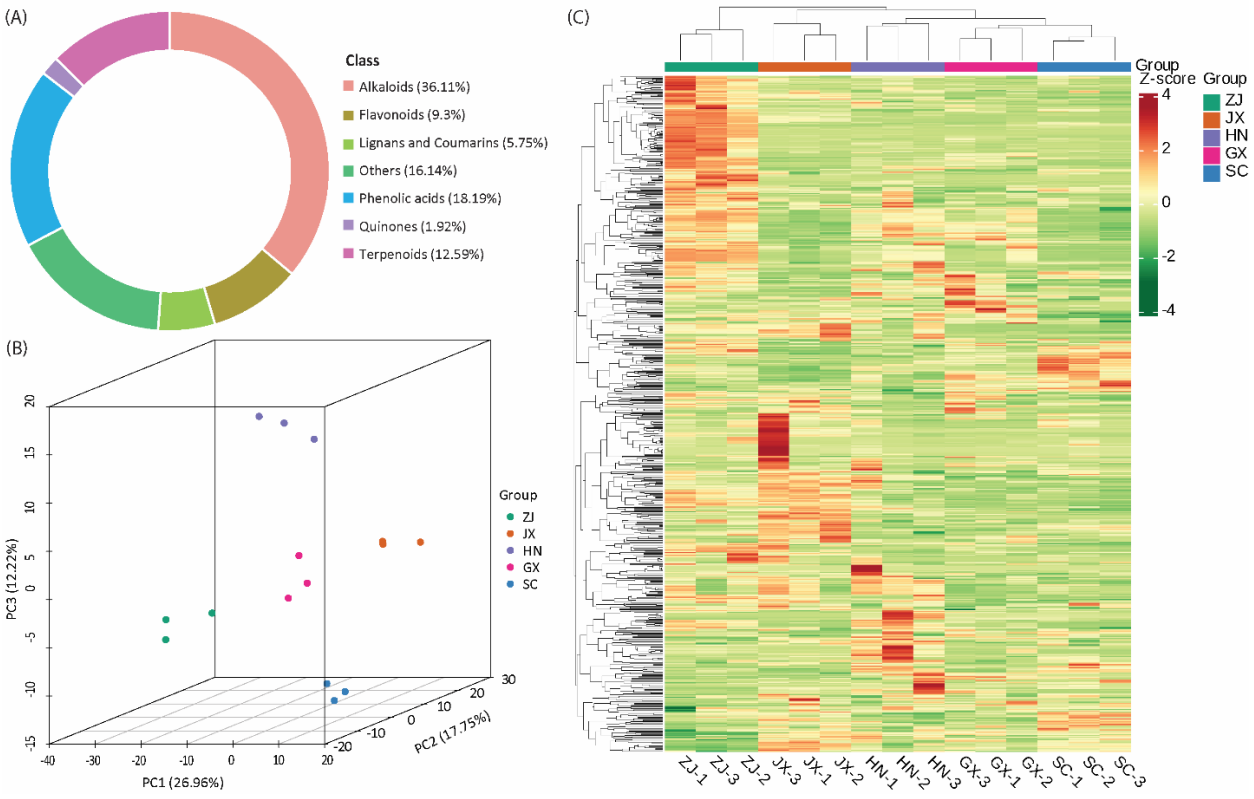


Fig. 2: *L. aurea* bulbs metabolomic analysis. (A) Ring diagram of metabolite class composition, with each color representing a metabolite category and the block area indicating the category's proportion. (B) 3D PCA of metabolites, with PC1, PC2, and PC3 representing the first, second, and third principal components, respectively, and percentages indicating each component's variance explained; each point represents a sample, with color denoting group membership. (C) Hierarchical clustered heatmap of metabolites across different experimental groups, with the sample names on the X-axis, metabolites on the Y-axis, and color gradients representing Z-scores for each data point.

Table 1: Summary of DAMs across different compared groups.

Group name	All significant differences	Down-regulated	Up-regulated
GX_vs_HN	171	120	51
GX_vs_JX	238	161	77
GX_vs_ZJ	283	211	72
HN_vs_JX	229	115	114
HN_vs_ZJ	216	160	56
JX_vs_ZJ	298	208	90
SC_vs_GX	203	104	99
SC_vs_HN	211	146	65
SC_vs_JX	226	156	70
SC_vs_ZJ	319	244	75

The fold change (FC) values of metabolites in the comparison group were calculated to highlight metabolic differences, with the top 10 up-regulated and down-regulated metabolites presented in a dynamic distribution diagram (Fig. S2). FC values across groups were compared, and bar charts of the top 20 metabolites with the highest FC between groups indicated that *L. aurea* bulb metabolites decrease with lower altitude and latitude but increase with higher longitude (Fig. S3). A clustering heatmap of differential metabolites revealed that samples from lower altitude/latitude or higher longitude (HN, JX, ZJ) exhibited higher Z-scores in Alkaloids, Terpenoids, and Flavonoids compared to SC or GX (Fig. S4). GX vs JX showed higher Z-scores in Phenolic acids, Flavonoids, Lignans, and Coumarins.

UV scaling followed by K-means cluster analysis showed the trend of the relative content of differential metabolites in different groups in two subclasses. In subclass 1, 259 metabolites were ranked in the following order of abundance ZJ>HN>GX>JX>SC, clearly influenced by the altitude (Fig. S5A). In subclass 2, however, the order for the remaining 721

metabolites was JX>HN>SC>GX>ZJ (Fig. S5B). This result showed that geographic location and altitude do not necessarily predict cluster similarity, indicating that local environmental factors may outweigh simple spatial distance in shaping phenotypic or molecular traits. However, the Venn diagram analysis of the DAM showed the highest number of common DAM when ZJ was compared to others (Fig. S6D), but the lowest number of common DAM was when GX was compared to others (Fig. S6A).

The KEGG analysis revealed several significantly enriched ($p<0.05$) DAM-associated metabolic pathways across the compared groups. However, after multiple comparison corrections, only the Flavonoid Biosynthesis (KO00941) pathway ($P=0.017$, cluster frequency 42.86%) in the GX_vs_HN comparison and the Phenylpropanoid Biosynthesis (KO00940) pathway ($P=0.028$, cluster frequency 25%) in the SC_vs_JX comparison remained significant (Table S3).

3.2. Spatial Influence on *L. aurea* Bulbs Gene Expression

Transcriptome sequencing of 15 samples from 5 locations generated 103.42 Gb of clean data, with each sample producing at least 6 Gb of clean reads and Q30 base percentages over 93%, indicating high sequencing quality. The assembly yielded a database of 189,456 unigenes with an average length of 1,060 bp, N50 of 1,478 bp, and N90 of 488 bp. Alignment with KEGG, NR, Swiss-Prot, GO, COG/KOG, and Trembl identified 77,301, 102,320, 73,923, 88,317, 62,084, and 101,653 homologs, covering 40.8%, 54.01%, 39.02%, 46.62%, 32.77%, and 53.66% of the sequences, respectively (Fig. 4A). Pfam comparison revealed 54,940 homologs with 29% coverage, and NR BLAST hits showed a 48.06% match with *Asparagus officinalis* (Fig. 4B). GO classification revealed enriched terms in biological process (cellular process, metabolic process, response to stimulus), cellular component (cellular anatomical entity, protein-containing complex), and molecular function [binding, catalytic activity, ATP (Adenosine triphosphate)-dependent activity] (Fig. 4C). KOG analysis categorized unigenes into 25 functional classes, with the largest group being "general function prediction only" (R, 14,692 unigenes), followed by "translation, ribosomal structure and biogenesis" (J, 7,163 unigenes), and "function unknown" (S, 5,536 unigenes), along with categories such as "posttranslational modification" (O), "transcription" (K), and "replication, recombination and repair" (L) (Fig. 4D).

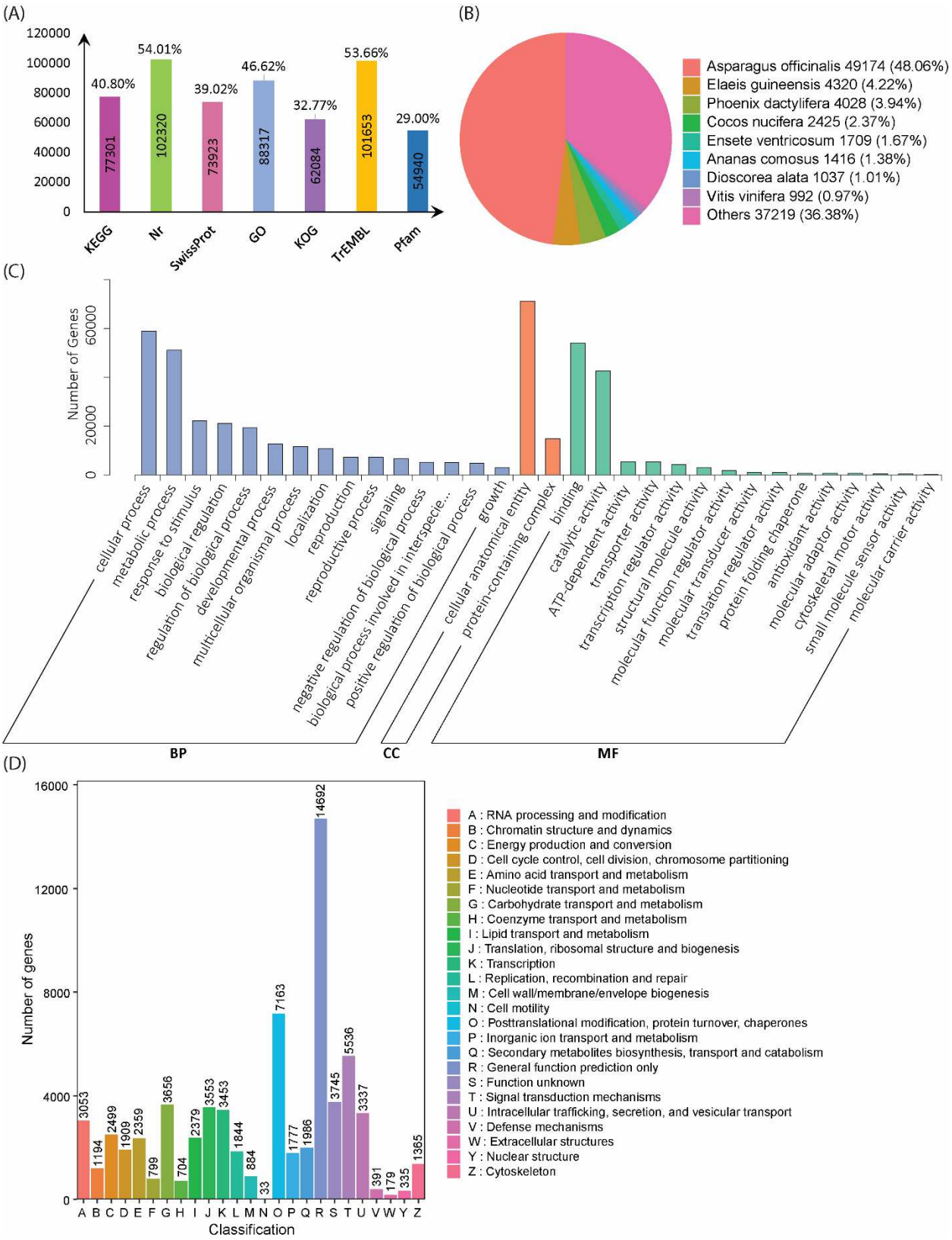


Fig. 3: Function annotations of transcriptome sequencing. (A) Functional annotation numbers of unigenes in the KEGG, NR, Swiss-Prot, GO, COG/KOG, Trembl, and Pfam

databases. (B) Annotated species distribution in the NR database. (C) GO classification histogram, where the horizontal axis represents secondary GO entries, and the vertical axis represents the number of genes annotated in each GO entry. (D) KOG classification chart, where the horizontal axis represents functional classification codes of KOG IDs, and the vertical axis represents the number of genes in each category. Different colors indicate different classifications, and the legend provides the code with its functional description.

FPKM analysis revealed 60600 DEGs with $|\log_2\text{Fold Change}| >= 1$ and FDR < 0.05 (Table 2) (Table S4). There was a significant difference in up-and down-regulated genes between groups, suggesting a location-specific difference in the *L. aurea* gene expression.

Table 2: Summary of DEGs across different compared groups.

Group name	All significant differences	Down-regulated	Up-regulated
GX_vs_HN	11253	5619	5634
GX_vs_JX	23871	12569	11302
GX_vs_ZJ	20792	9468	11324
HN_vs_JX	22693	11686	11007
HN_vs_ZJ	20763	9288	11475
JX_vs_ZJ	27933	12232	15701
SC_vs_GX	15376	8337	7039
SC_vs_HN	17867	9593	8274
SC_vs_JX	27846	15503	12343
SC_vs_ZJ	24879	12166	12713

The FPKM distribution box plot showed consistent gene expression with narrow IQRs in the ZJ, JX, HN, and GX groups (Fig. S7A). In contrast, the SC group displayed greater variability and a wider IQR, indicating gene expression fluctuation. This was further confirmed by the FPKM density distribution (Fig. S7B) and violin plot (Fig. S7C). Pearson correlation analysis (Fig. 4A) revealed strong consistency in gene expression within most groups, especially ZJ, JX, HN, and GX, while SC showed weaker correlations, indicating higher variability. 3D PCA analysis (Fig. 4B) showed that PC1 (19.41%), PC2 (15.09%), and PC3 (12.76%) captured significant variance, with distinct clustering of ZJ and JX, overlap in SC and GX, and separation of the HN group.

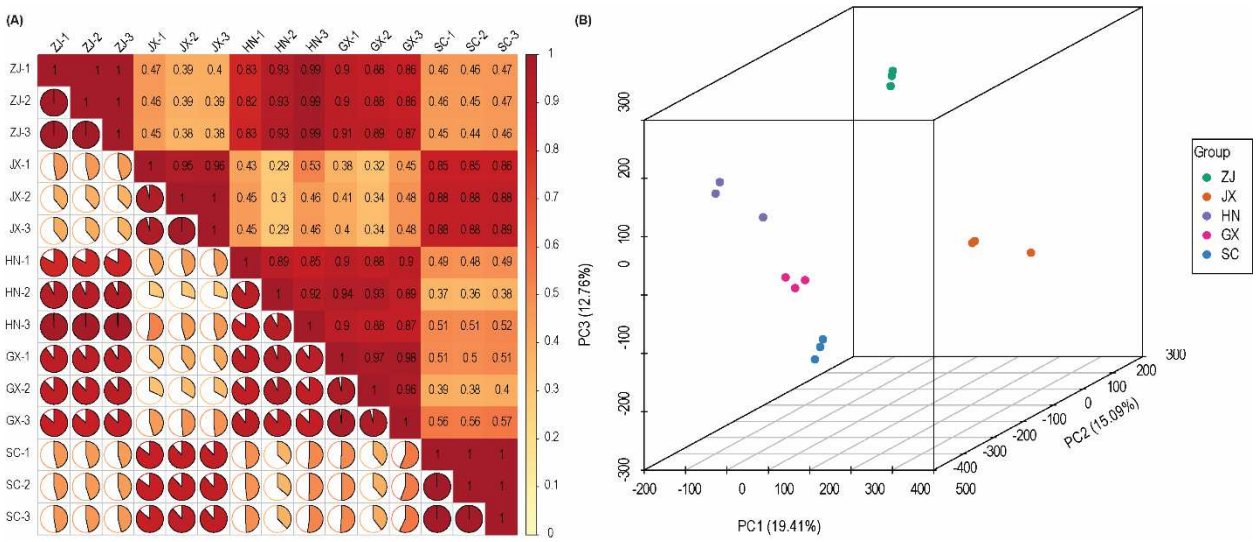


Fig. 4: *L. aurea* bulbs Gene Expression Pattern. (A) The sample correlation plot, where Pearson's Correlation Coefficient (r) is used to evaluate the correlation between biological replicates. The closer the absolute value of r is to 1 (the redder the color), the stronger the correlation between the two replicate samples. (B) 3D PCA plot of gene expression where PC1, PC2, and PC3 represent the first, second, and third principal components, and the percentage represents the explanation rate of the principal component for the data set; each point in the figure represents a sample, and the same color represents samples in the same group.

The volcano plots (Fig. S8) revealed that most clusters were non-regulated across comparisons, with several key clusters showing significant upregulation or downregulation, particularly in the order JX_vs_ZJ > SC_vs_JX > SC_vs_ZJ > GX_vs_JX. The highest number of significantly up-regulated genes was observed in the JX_vs_ZJ comparison (15,701). The SC group exhibited a higher proportion of non-regulated genes, especially in the SC_vs_GX (45,224) and SC_vs_ZJ (35,721) comparisons. The radar chart (Fig. S9) identified clusters with substantial fold changes between comparisons. In the JX_vs_ZJ comparison, clusters such as 92568.3 and 90745.2 showed significant fold changes, while in SC_vs_JX, clusters 90372.1 and 63903.15 were notable. Similarly, in SC_vs_ZJ, clusters 50696.3 and 93131.15 exhibited large fold changes, and in GX_vs_JX, clusters 93131.19 and 63903.15 were significantly changed. Notably, cluster 90372.1 appeared in six comparisons (GX_vs_JX, GX_vs_ZJ, HN_vs_JX, HN_vs_ZJ, SC_vs_JX, and SC_vs_ZJ), and cluster 51222.2 was found in four comparisons (GX_vs_JX, GX_vs_ZJ, SC_vs_JX, and SC_vs_ZJ). Several other top clusters were also common across two or three comparisons.

The heatmap color bar and hierarchical clustering dendrogram between groups (Fig. S10) showed distinct gene expression patterns between *L. aurea* bulbs from different locations, with clear group separation. Some genes varied within the same group, suggesting environmental or genetic influences. Clustering revealed that SC was most closely related to GX, followed by ZJ, HN, and JX (Fig. 5).

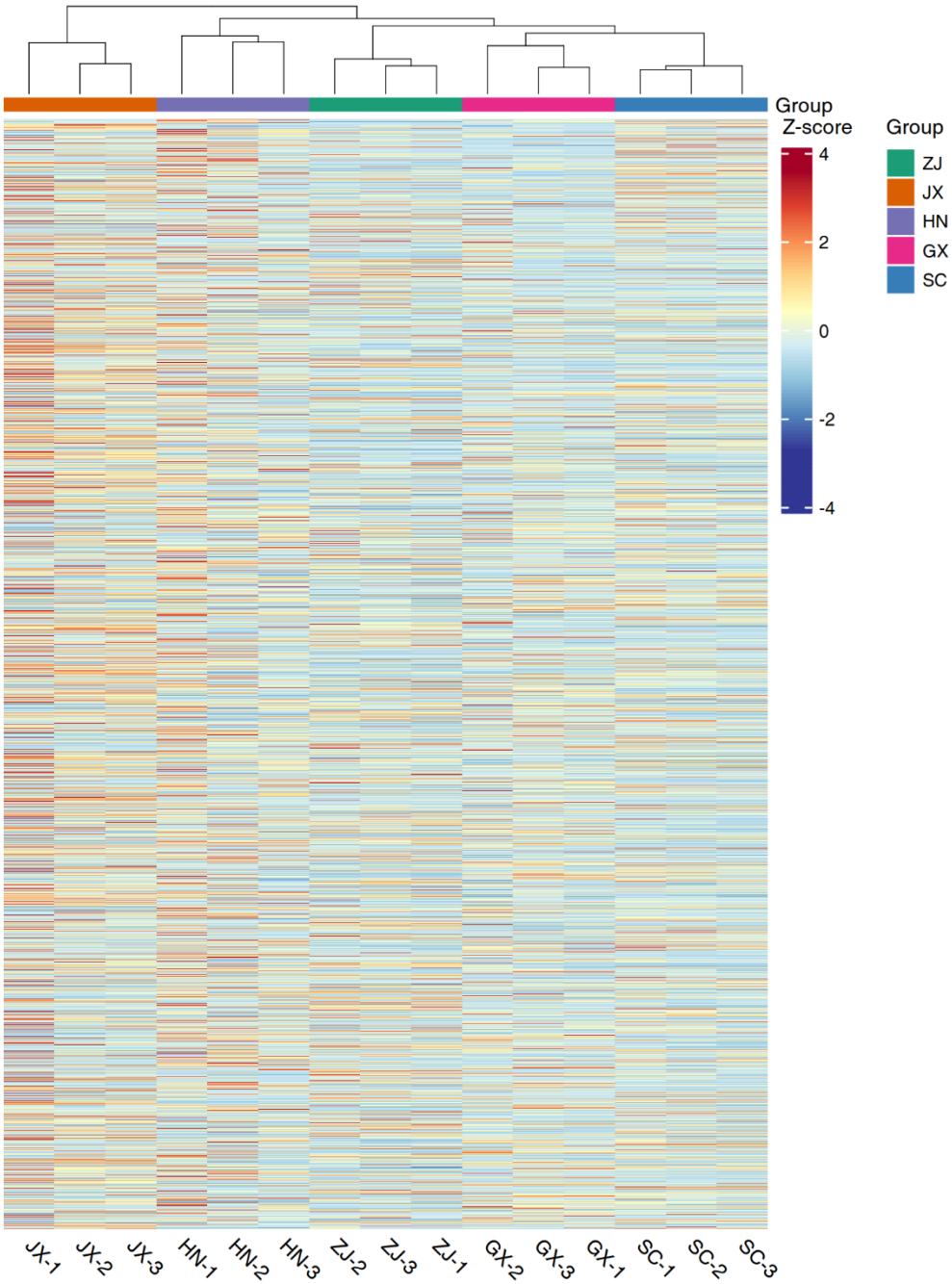


Fig. 5: Differential gene clustering heat map of *L. aurea* bulbs from different locations. Here, the horizontal axis represents the sample name and hierarchical clustering results, and the vertical axis represents the differentially expressed genes and hierarchical clustering results. Red represents high expression, and blue represents low expression.

K-means cluster analysis of 60,600 DEGs revealed distinct gene expression patterns across 10 subclasses (Fig. S11). In subclass 1, 7,005 DEGs were ranked in the order ZJ>JX>HN>GX>SC (Fig. S11A), while in subclass 2, 8,501 DEGs were ranked SC>GX>HN>JX>ZJ (Fig. S11B), indicating altitude influence. Sharp peaks in subclasses 3,

5, and 6 (Fig. S11C, E, F) were observed in HN; ZJ and SC in subclass 4 (Fig. S11D); GX in subclasses 7 and 8 (Fig. S11G, H); and JX in subclasses 9 and 10 (Fig. S11I, J), suggesting environmental or other factors. The Venn diagram showed the highest number of unique DEGs in JX compared to other groups (Fig. S12C), while GX had the fewest unique DEGs (Fig. S12A).

KEGG analysis of DEGs identified several significant DEG-associated pathways ($p<0.05$) across the compared groups, with 10 pathways remaining significant after P-value adjustment (Table S5). The Protein processing in endoplasmic reticulum pathway (KO04141) was significant in SC_vs_JX (P = 0.0000003901), HN_vs_ZJ (P = 0.0000008429), GX_vs_HN (P = 0.0000074452), SC_vs_HN (P = 0.0000490227), JX_vs_ZJ (P = 0.0000816909), HN_vs_JX (P = 0.0002185169), GX_vs_JX (P = 0.0085480832), and SC_vs_ZJ (P = 0.0091569183). The Ribosome pathway (KO03010) was significant in SC_vs_HN (P = 0.0000039810) and HN_vs_ZJ (P = 0.0001579419). The Starch and sucrose metabolism pathway (KO00500) was significant in SC_vs_GX (P = 0.0067584993) and GX_vs_JX (P = 0.0113422941). The Cutin, suberine, and wax biosynthesis pathway (KO00073) was significant in SC_vs_GX (P = 0.0004739915) and SC_vs_ZJ (P = 0.0195882035). Other significant pathways included the Spliceosome pathway (KO03040) in HN_vs_ZJ (P = 0.0000519557), Biosynthesis of secondary metabolites (KO01110) in SC_vs_ZJ (P = 0.0096887642), Biosynthesis of unsaturated fatty acids (KO01040) in SC_vs_HN (P = 0.0066542895), Phenylpropanoid biosynthesis (KO00940) in JX_vs_ZJ (P = 0.0037528038), Linoleic acid metabolism (KO00591) in GX_vs_HN (P = 0.0002745666), and Glutathione metabolism (KO00480) in GX_vs_ZJ (P = 0.0199734744).

3.3. Integrated Analysis

Combined PCA analysis of the transcriptome and metabolome showed distinct clustering of samples, particularly from ZJ and SC (Fig. 6A), with clear separation between regions, indicating unique gene expression profiles. Metabolome data (Fig. 6B) also revealed differentiation, with ZJ displaying a unique metabolic profile, while HN and GX clustered closely together. JX samples showed overlapping and distinct separation patterns, suggesting varying metabolic signatures. Overall, ZJ exhibited the most pronounced differences in both transcriptomic and metabolomic profiles.

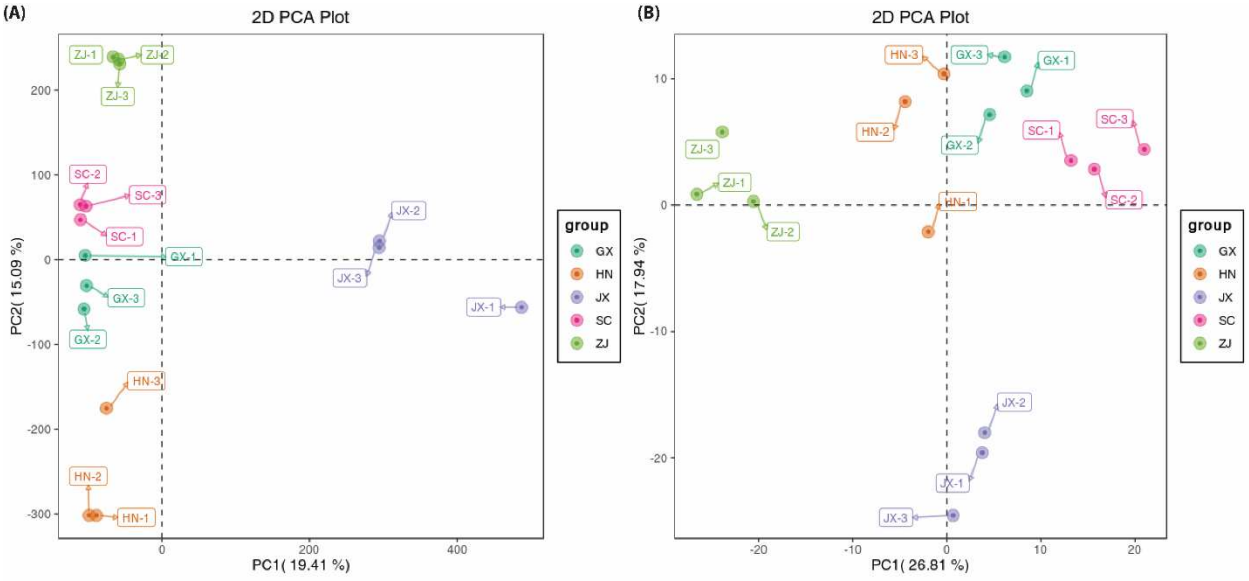


Fig. 6: 2D PCA analysis of *L. aurea* transcriptomes (A) and metabolomes (B). The horizontal axis represents principal component 1, the vertical axis represents principal component 2, and points of different colors represent samples in different groups.

A pearson correlation between the environmental variables and the K-means cluster analysis sub-classes of the DAMs and DEGs demonstrated a complex interplay between them (Fig. 7). Gene subclass 6 showed high ($p < 0.001$) positive correlation with Se concentration whereas metabolite subclass 1 showed moderate positive ($p < 0.01$) correlation with soil pH and week positive correlation ($p < 0.05$) with elevation. Gene subclass 1 showed a weak negative correlation ($p < 0.05$) with the highest temperature and longitude. The dendrogram showed that gene subclass 6 and metabolite subclass 2 are highly correlated and grouped. Metabolite subclass 1 and gene subclass 9 are highly correlated and grouped, which in turn are related to gene subclass 1.

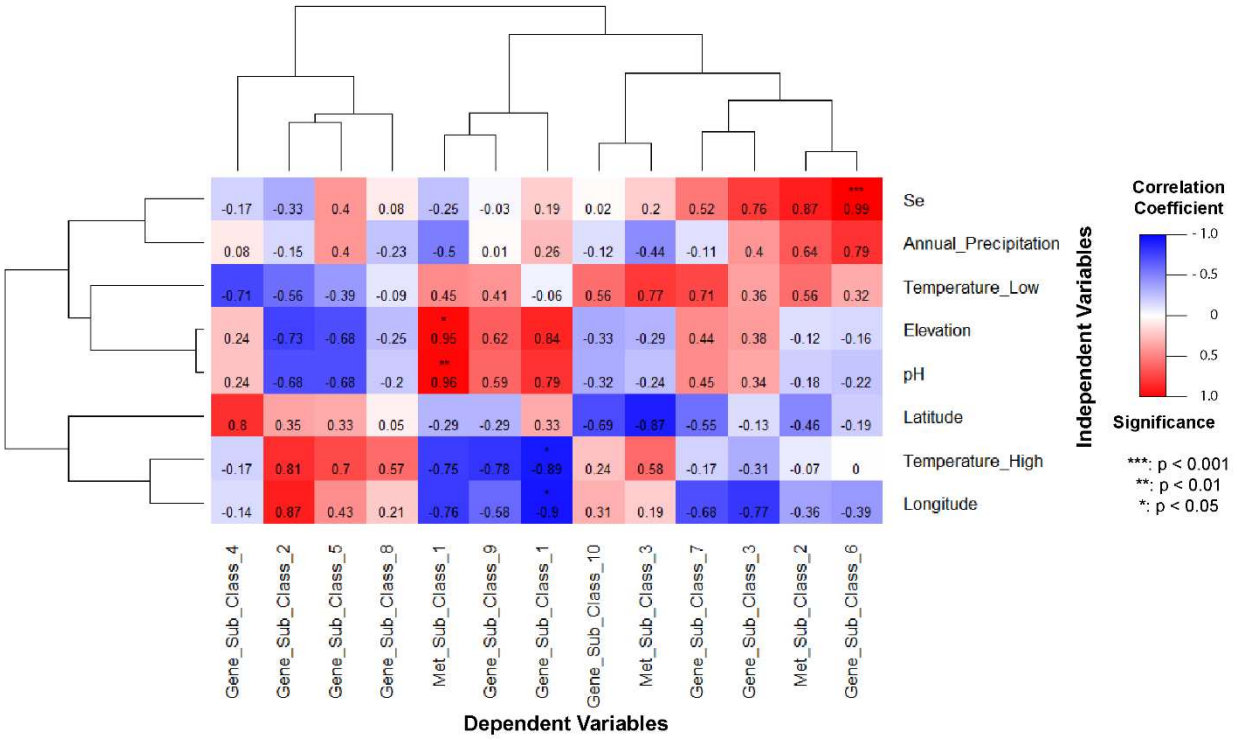


Fig. 7: Correlation heatmap with dendrogram analysis between environmental and spatial variables vs metabolites and transcriptome k-means analysis. Here, the color from blue to red indicates negative to positive correlation coefficients, and the asterisk indicates a significance label.

Combined KEGG pathway analysis revealed that the top pathways containing more than five DAMs were: Metabolic pathways (KO01100), Biosynthesis of secondary metabolites (KO01110), Phenylpropanoid biosynthesis (KO00940), Tryptophan metabolism (KO00380), and Flavonoid biosynthesis (KO00941) (Fig. 8A, Table S6). The top pathways containing more than 100 DEGs were: Metabolic pathways (KO01100), Biosynthesis of secondary metabolites (KO01110), Biosynthesis of amino acids (KO01230), Biosynthesis of cofactors (KO01240), Phenylpropanoid biosynthesis (KO00940), Glycerophospholipid metabolism (KO00564), Tryptophan metabolism (KO00380), and Glutathione metabolism (KO00480) (Fig. 8A, Table S6). After enrichment analysis of DAMs, six KEGG pathways were identified; following P-value adjustment, only Flavonoid biosynthesis (KO00941) and Phenylpropanoid biosynthesis

(KO00940) remained significant in GX_vs_HN and SC_vs_JX, respectively (Fig. 8B, Table S7). KEGG enrichment analysis of DEGs revealed twelve KEGG pathways, with only Phenylpropanoid biosynthesis (KO00940) and Biosynthesis of secondary metabolites (KO01110) remaining significant in JX_vs_ZJ and SC_vs_ZJ, respectively (Fig. 8B, Table S7).

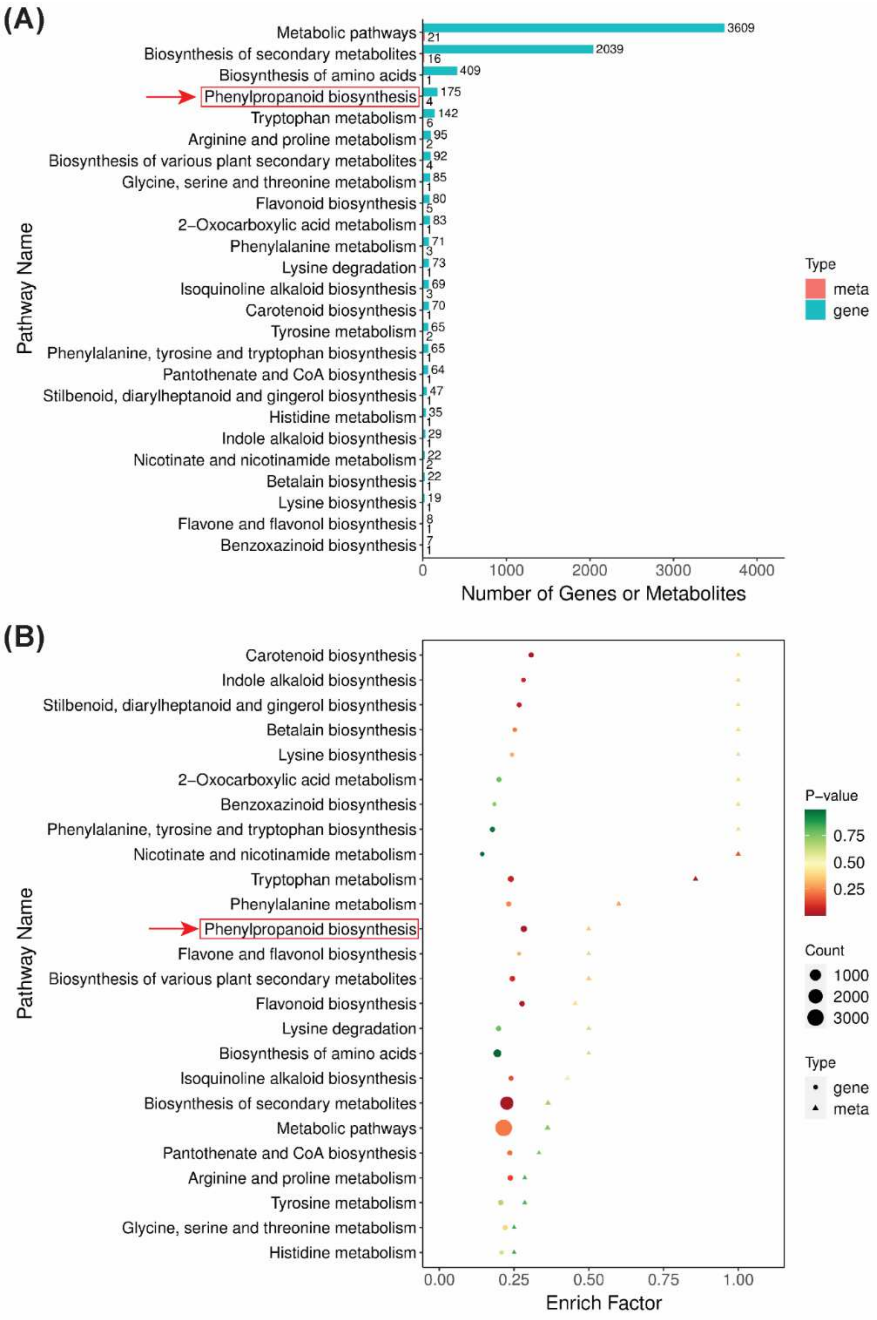


Fig. 8: *L. aurea* transcriptome and metabolome combined KEGG Enrichment Analysis between JX_vs_ZJ Groups.

The bar chart (A) displays the 25 pathways with the highest P-values in multi-omics analysis, where the bar length represents the number of differential metabolites and differential genes enriched in each pathway. The bubble diagram (B) shows the top 25

pathways with the highest P-values, where the X axis represents the enrichment factor, the bubble shape indicates DEGs or DAMs, with their size representing the number of genes or metabolites, and the color denotes the P-value. To analyze the regulatory network of DAMs, correlation analysis between DEGs and DAMs was performed in each group, revealing 67,962 DEGs significantly correlated with 732 DAMs (Table S8). Expression trend analysis showed that multiple metabolites were positively or negatively regulated by several genes (Fig. S13). For example, in GX_vs_HN (Fig. S13A), positive correlations between genes and metabolites were fewer compared to SC_vs_JX (Fig. S13B), JX_vs_ZJ (Fig. S13C), and SC_vs_ZJ (Fig. S13D). KGML analysis across groups revealed that the ath00010 pathway was the most highly clustered, followed by ath00260 and ath00030 in SC_vs_JX (Fig. S14B) and SC_vs_ZJ (Fig. S14D). In GX_vs_HN (Fig. S14A), the second most clustered pathway was ath00020, while in JX_vs_ZJ (Fig. S14C), it was ath00520. The ath00260 pathway was the most upregulated cluster in SC_vs_JX, JX_vs_ZJ, and SC_vs_ZJ. CCA analysis of genes and metabolites related to the KO00941 pathway in GX_vs_HN (Fig. S15A), KO00940 pathway in SC_vs_JX (Fig. S15B), KO01110 pathway in JX_vs_ZJ (Fig. S15C), and KO00940 pathway in SC_vs_ZJ (Fig. S15D) showed strong correlations. In GX_vs_HN, MWSHY0098 (Epigallocatechin), MWSHY0037 (Isoliquiritigenin*), and others were down-regulated, while MWS0178 [Chlorogenic acid (3-O-Caffeoylquinic acid)*] was up-regulated. In SC_vs_JX, MWS2212 (Caffeic acid), MWS0906 (Coniferin), and HJN003 (1-O-Sinapoyl-β-D-glucose) were up-regulated, while MWS2208 (Ferulic acid) and others were down-regulated. In SC_vs_ZJ, MWS2212, MWSHY0037, and MWS0178 were up-regulated, and other metabolites like MWS0677 (N-Acetyl-5-hydroxytryptamine) were down-regulated. MWS0178 was up-regulated in all but SC_vs_JX, indicating a positive correlation with latitude, while MWS2212 was inversely correlated with latitude. Gene clusters related to their regulation were identified. MWSHY0098 and MWSCX015 (Caffeic aldehyde) were influenced by longitude. Ten metabolites, such as LSKP211262 (Secoisolariciresinol), decreased with lower elevation, and four metabolites, including MW0139629 (Sakuranetin), decreased with increasing latitude and decreasing altitude, suggesting that latitude, altitude, and environmental factors affect *L. aurea* gene expression and metabolism.

4. Discussion

This integrated, multi-omics investigation sheds insight into the geographical and environmental factors that depend on the spatial variation of *L. aurea* (L'Hér.) bulb metabolites and transcriptomes, which have both pharmacological and ecological implications. The metabolomic data revealed considerable differences in the composition of secondary metabolites, including alkaloids, flavonoids, phenolic acids, and terpenoids, among the various samples. The transcriptomic data further confirmed this, revealing location-specific variations in gene expression related to these metabolic pathways. The integration of both data sets highlighted key genes and metabolic pathways that are responsive to environmental cues. For instance, *L. aurea* bulbs from higher altitudes showed an increase in the abundance of flavonoids and terpenoids, which are known to play vital roles in plant defense mechanisms against environmental stress. The transcriptome data indicated upregulation of genes involved in the biosynthesis of these compounds, reinforcing the idea that plants modulate both gene expression and metabolite accumulation to adapt to their environment. Environmental parameters such as temperature, precipitation, soil pH, and Se concentration are also known to affect plant metabolism [25, 26, 32, 33]. In their study M Quan*, et al.* [12] found humus soil with looser texture and lower moisture is more

suitable for artificial cultivation of *L. aurea* for increasing photosynthetic rate, biomass, and lycorine content. More detailed observation by J Liang*, et al.* [21] revealed that adequate irrigation is required during vegetative growth, but mild water deficit increases alkaloid content at a later growth stage. Although in our study, annual precipitation showed no correlation with metabolite or gene subclasses; soil Se, pH, and elevation demonstrated positive correlation with different subclasses of metabolites and genes (Fig. 7). Our result also coincides with Y-W Zuo*, et al.* [3] findings of positive correlation between lycorine concentration with soil pH, water content and Se levels which inturn positively influence soil bacterial populations. Similarly, M Quan*, et al.* [12] also showed a positive correlation between soil pH and lycorine from *L. aurea*. Se is an essential trace element known for its role in modulating plant growth and metabolism [48]. The highly positive correlation it showed with gene subclass 6 suggests its potential involvement in gene expression regulation and related metabolic pathways, such as the phenylpropanoid biosynthesis pathway (Fig. 8). G Guo*, et al.* [48] reported Se to positively influence anthocyanin biosynthesis, which is a subsequent product of the phenylpropanoid biosynthesis pathway we found. Similarly, Y-W Zuo*, et al.* [3] also deduced that Se positively influences phenylpropanoid production by reducing oxidative stress. We found gene subclass 1 was significantly negatively correlated with the highest temperature and longitude (Fig. 7), which was also observed in the paired FC bar chart of metabolites, where upregulated metabolites decreased in descending altitude (Fig. S3). Latitude and altitude were reported to cause metabolic differences in various plants, resulting in differential expression of genes, metabolites, and triggering defense mechanisms [23, 24, 28, 31]. The effect of longitude on genes and metabolites was also apparent in the pairwise heat map, where higher longitude showed higher Z-score in secondary metabolites like Alkaloids, Terpenoids, Flavonoids, Phenolic acids, Lignans, and Coumarins (Fig. S4). Longitude can influence the intensity and duration of sunlight exposure, which in turn may affect the production of these compounds, as many of these secondary metabolites were reported to accumulate in high concentration under lower light exposure [49]. However, k-means cluster analysis showed altitude together with other environmental factors might influence *L. aurea* metabolites (Fig. S5) and gene expression (Fig. S11). This was also confirmed from the Venn diagram analysis of metabolite (Fig. S6) and transcript (Fig. S12) compared groups, as it showed the number of unique DAMs and DEGs increases with an increase in altitude and latitude.

KEGG pathway analysis of DAMs and DEGs revealed Metabolic pathways (KO01100), Biosynthesis of secondary metabolites (KO01110), Phenylpropanoid biosynthesis (KO00940), and Tryptophan metabolism (KO00380) were commonly enriched (Fig. 8, Table S6 and S7), of which, after p-adjustment, only Phenylpropanoid biosynthesis (KO00940) remained significant. A previous study reported that Se causes differential expression of Phenylpropanoid biosynthesis (KO00940) in *Zea mays* L., affecting anthocyanin biosynthesis [48]. Correlation analysis found 67,962 DEGs significantly associated with 732 DAMs (Table S8) with diverse positive and negative correlations between genes and metabolites across groups (Fig. S13), further demonstrating the influence of environmental factors. Also, the CCA analysis (Fig. S15) again showed pathways such as KO00941, KO00940, KO01110, and KO00940 (in GX_vs_HN, SC_vs_JX, JX_vs_ZJ, and SC_vs_ZJ, respectively) were among the pathways with the strongest correlation. Among the significant metabolites in these pathways, HJN003 (1-O-Sinapoyl-β-D-glucose), MWS2212 (Caffeic acid), and MWSHY0037 (Isoliquiritigenin*) were inversely related, but MWS0178 [Chlorogenic acid (3-O-Caffeoylquinic acid)*] was positively associated with the latitude. Whereas MWSHY0098

(Epigallocatechin) and MWSCX015 (Caffeic aldehyde) were positively associated with longitude. Ten metabolites, including LSKP211262 (Secoisolariciresinol), dropped with lower height, while four metabolites, including MW0139629 (Sakuranetin), declined with rising latitude and decreasing altitude.

1-O-Sinapoyl-β-D-glucose, a glucosyl hydroxycinnamic acid, plays a crucial role in plant metabolism and defense in plants [50, 51]. Caffeic acid is essential for lignin synthesis, as well as turgor pressure, water flux, phototropism, cell expansion, and growth regulation [52]. Isoliquiritigenin, a flavonoid biologic from the licorice (*Glycyrrhiza uralensis*) root, is a widely used food and remedy with biological properties including anti-inflammatory, antioxidant, neuroprotective, and anticancer activity against several types of cancers [53]. Chlorogenic acid (3-O-Caffeoylquinic acid)*, the ester of Caffeic acid, is also involved in plant defense (against pathogen, herbivores), stresses (such as UV, heavy metal toxicity, oxidative stress), plays a role in lignin synthesis for wound healing, and has anticancer properties [54]. Other significant metabolites involved in plant defense and lignin synthesis include Epigallocatechin [55], Caffeic aldehyde [56], and Secoisolariciresinol [57], whereas Sakuranetin was primarily involved as a phytoalexin and protection against biotic and abiotic stresses [58]. These findings suggest that latitude, altitude, and environmental factors concurrently affect *L. aurea* gene expression and metabolism, many of which have known therapeutic use. Similar spatial effect on *Lavandula angustifolia* Mill. metabolite outcome was also observed by S Demasi*, et al.* [31]. Our investigation, however, showed the importance of studying the spatial and environmental factors for improving *L. aurea* biologics.

5. Conclusions

The integration of metabolomics and transcriptomics in this study has provided a holistic understanding of the spatial variations in the biochemical and genetic profiles of L. aurea bulbs. Environmental factors such as soil pH and selenium availability were shown to influence both gene expression and metabolite accumulation, highlighting the adaptive nature of *L. aurea* to its environment. Further analysis revealed that spatial factors, such as latitude and altitude, influence metabolic outcome, leading to this adaptation. We reported combinatorial, spatial, and ecological effects on *L. aurea* through a multi-omics study, implicating the necessity of this study for its cultivation and medicinal value. This multi-omics approach not only enhances our understanding of plant-environment interactions but also offers valuable insights into optimizing cultivation practices and improving the quality of medicinal plants. Future studies could focus on further elucidating the specific regulatory networks that link gene expression with metabolite biosynthesis, particularly in response to environmental cues.

6. Abbreviations

Abbreviation	Full Form
ALDH	Aldehyde Dehydrogenase
BLAST	Basic Local Alignment Search Tool
CCA	Clusters of Orthologous Groups
CDS	Potential Coding Regions
COG	Orthologous Gene Families in Prokaryotes
DAMs	Differentially Accumulated Metabolites
DEGs	Differentially Expressed Genes

DESeq2	Differential Gene Expression Analysis Based on the Negative Binomial Distribution
ESI	Electrospray Ionization
FDR	False Discovery Rate
FPKM	Fragments Per Kilobase of Transcript Per Million Fragments Mapped
GO	Gene Ontology
HCA	Hierarchical Cluster Analysis
IQR	Interquartile Range
KEGG	Kyoto Encyclopedia of Genes and Genomes
KGML	Kegg Markup Language
KOG	Eukaryotic Orthologous Groups
MISA	Microsatellite Identification Tool
MRM	Multiple Reaction Monitoring
MS/MS	Tandem Mass Spectrometry
MSEA	Metabolite Set Enrichment Analysis
NCBI	National Center for Biotechnology Information
NR	NCBI Non-Redundant Protein Sequences Database
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PCA	Principal Component Analysis
PCC	Pearson Correlation Coefficients
PPIs	Predicted Protein-Protein Interactions
QQQ	Triple Quadrupole Mass Spectrometry
SSR	Simple Sequence Repeats
STRING	Functional Protein Association Networks
TIC	Total Ion Chromatogram
TrEMBL	Translated EMBL Nucleotide Sequence Database
UPLC	Ultra-Performance Liquid Chromatography
VIP	Variable Importance Plots
WGCNA	Weighted Correlation Network Analysis

7. Supplementary Information

7.1. Supplementary Figures

Fig. S1: OPLS-DA score plot showing L. aurea bulb samples from different locations. Panels (A)–(J) depict pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, (J) SC_vs_ZJ. The horizontal axis represents the predicted principal component (variation between groups), and the vertical axis represents the orthogonal principal component (variation within groups). Percentages indicate the variance explained by each component. Points represent samples, with colors denoting group membership.

Fig. S2: Dynamic distribution diagram of L. aurea metabolite content differences. Panels (A)–(J) show pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, and (J) SC_vs_ZJ. The horizontal axis represents the cumulative number of substances arranged by difference multiple, while the vertical axis shows the logarithm of the difference multiple (base 2). Each

point represents a substance, with green points indicating the top 10 downgraded substances and red points indicating the top 10 upgraded substances.

Fig. S3: Fold difference bar chart of L. aurea bulb metabolites. Panels (A)–(J) show pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, and (J) SC_vs_ZJ. The horizontal axis represents the $\log_2$ of differential metabolite fold change (FC), and the vertical axis represents the differential metabolites. Red bars indicate upregulated metabolites, while green bars indicate downregulated metabolites.

Fig. S4: Heat map of L. aurea bulb metabolite classes. Panels (A)–(J) show pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, and (J) SC_vs_ZJ. The horizontal axis represents sample names, while the vertical axis shows differential metabolite information. The Group indicates sample grouping and colors represent relative metabolite content after standardization (red for high content, green for low content). Class refers to the first-level classification of substances.

Fig. S5: K-Means plot of L. aurea bulbs differential metabolites. This analysis identified two distinct clusters, with the standardized scores for clusters 1 and 2 presented in panels (A and B). The horizontal axis represents the sample grouping, the vertical axis represents the standardized relative content of metabolites, Sub class represents the metabolite category number with the same change trend.

Fig. S6: Venn diagram of differences among groups of L. aurea bulbs metabolites. Panel A-E shows ven diagram between different compared groups: (A) GX_vs_ZJ, GX_vs_HN, SC_vs_GX, and GX_vs_JX; (B) HN_vs_ZJ, GX_vs_HN, SC_vs_HN, and HN_vs_JX; (C) JX_vs_ZJ, GX_vs_JX, SC_vs_JX, and HN_vs_JX; (D) JX_vs_ZJ, JX_vs_ZJ, SC_vs_ZJ, and HN_vs_ZJ; (E) SC_vs_ZJ, SC_vs_HN, SC_vs_GX, and SC_vs_JX. Here, each circle in the figure represents a comparison group. The numbers in the overlapping part of the circles represent the number of common differential metabolites between the comparison groups, and the numbers without overlapping parts represent the number of unique differential metabolites in the comparison groups.

Fig. S7: Quantitation of L. aurea bulbs Gene Expression. (A) The box plot of expression where the horizontal axis in the figure represents different samples; the vertical axis represents the logarithmic value of the sample expression FPKM. This figure measures the expression level of each sample from the perspective of the overall dispersion of the expression. (B) The expression density distribution diagram where the curves of different colors in the figure represent different samples. The horizontal axis of the points on the curve represents the logarithmic value of the corresponding sample FPKM, and the vertical axis of the points represents the probability density. (C) The violin plot of expression where different colors in the figure represent different samples, and the width of each violin graph reflects the number of transcripts at that expression level.

Fig. S8: Volcano plot of DEGs of L. aurea bulbs from different locations. Panels (A)–(J) show pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, and (J) SC_vs_ZJ. Here, the horizontal axis represents the fold change of gene expression, and the vertical axis represents the significance level of differentially expressed genes. Red dots represent

upregulated differentially expressed genes, green dots represent down-regulated differentially expressed genes, and gray dots represent non-differentially expressed genes.

Fig. S9: Radar charts depicting differentially expressed genes (DEGs) of L. aurea bulbs from different locations. Panels (A)–(J) show pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, and (J) SC_vs_ZJ. Each point represents a gene, and its position on the chart reflects the magnitude of the log2 fold change (log2FC) of the gene's expression across the two compared locations.

Fig. S10: Heatmap of differential gene clustering analysis for L. aurea bulbs from different locations. Panels (A)–(J) show pairwise comparisons: (A) GX_vs_HN, (B) GX_vs_JX, (C) GX_vs_ZJ, (D) HN_vs_JX, (E) HN_vs_ZJ, (F) JX_vs_ZJ, (G) SC_vs_GX, (H) SC_vs_HN, (I) SC_vs_JX, and (J) SC_vs_ZJ. The horizontal axis shows sample names and hierarchical clustering results, while the vertical axis represents differentially expressed genes with their respective hierarchical clustering results. Red indicates high gene expression, and blue indicates low gene expression.

Fig. S11: K-Means plot of L. aurea bulbs differential genes. Panels (A)–(J) show different subclass: (A) Subclass 1, (B) Subclass 2, (C) Subclass 3, (D) Subclass 4, (E) Subclass 5, (F) Subclass 6, (G) Subclass 7, (H) Subclass 8, (I) Subclass 9, and (J) Subclass 10. Here, the horizontal axis represents the sample, and the vertical axis represents the standardized expression level.

Fig. S12: Venn diagram of DEGs of L. aurea bulbs. Panels (A)–(E) show Ven diagram between different sample comparisons: (A) GX_vs_ZJ, GX_vs_HN, SC_vs_GX, and GX_vs_JX; (B) HN_vs_ZJ, GX_vs_HN, SC_vs_HN, and HN_vs_JX; (C) JX_vs_ZJ, GX_vs_JX, SC_vs_JX, and HN_vs_JX; (D) JX_vs_ZJ, GX_vs_JX, SC_vs_ZJ, and HN_vs_ZJ; (E) SC_vs_ZJ, SC_vs_HN, SC_vs_GX, and SC_vs_JX. Here, the non-overlapping area of the Venn diagram represents the differential genes unique to the differential grouping, and the overlapping area represents the differential genes shared by several overlapping differential groups.

Fig. S13: Expression trend analysis in GX_vs_HN (A), SC_vs_JX (B), JX_vs_ZJ (C), and SC_vs_ZJ (D). Here, the dots and boxes in the figure represent metabolites and genes, respectively. Red indicates upregulated genes/metabolites, green indicates down-regulated genes/metabolites, and blue indicates both upregulated and down-regulated genes.

Fig. S14: KGML analysis network diagram of the gene and the metabolites to the pathway between GX_vs_HN (A), SC_vs_JX (B), JX_vs_ZJ (C), and SC_vs_ZJ (D). The squares in the figure represent genes or gene products, circles represent metabolites, and diamonds represent pathway names. Red indicates that genes, gene products, or metabolites are upregulated, and green indicates that genes, gene products, or metabolites are downregulated.

Fig. S15: CCA analysis of the gene and the metabolites to the KO00941 pathway in GX_vs_HN (A), KO00940 pathway in SC_vs_JX (B), KO001110 pathway in JX_vs_ZJ (C), and KO00940 pathway in SC_vs_ZJ (D). The figure uses a cross to distinguish four regions. In the same region, the farther from the origin, the closer the distance, and the higher the correlation. Metabolites are marked purple, and genes are marked in red. If there are too many substances of a certain type, they will be displayed as dots to avoid text overlapping.

7.2.Supplementary Tables

Table S1: Metabolite quantity statistics

Table S2: Differentially Enriched Metabolites

Table S3: DEM KEGG summary

Table S4: Differentially Enriched Genes

Table S5: DEG KEGG summary

Table S6: Combined KEGG pathways

Table S7: Combined KEGG enrichment analysis

Table S8: Combined Correlation Analysis

8. Declarations

Funding

This work was funded by the National Natural Science Foundation of China (32070367), the Hunan Provincial Natural Science Foundation (2024JJ7367), and the earmarked fund for HARS.

Availability of data and materials

All acquired raw sequencing data from Lycoris aurea [Taxonomy ID: 152838] transcriptome and gene expression are submitted to the SRA at the NCBI database under accession number PRJNA1328393 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1328393) SUB14823235. We summarized the datasets used in this manuscript and presented them as supporting information for publication. The corresponding author will make any other relevant information available upon reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Clinical trial number

Not applicable.

Authors' contributions

Conceptualization and original draft, SL, LS, TZ; methodology, LS, TY; software and visualization, SL, TZ; formal analysis, TY; supervision, review & editing, LS, QM; project administration and funding acquisition, LS, QM

Acknowledgments

Not applicable.

Competing interests

The authors declare no conflict of interest.

9. References

1. Peng Y, Wei J, Yang L. The complete chloroplast genome of *Lycoris aurea* (L'Hér.) Herb. Mitochondrial DNA Part B Resources. 2020; 5:788-789. https://doi.org/10.1080/23802359.2020.1715296.

2. Wang Z, Shu X, Wang N, Cheng G, Zhang F. 'E Huang Xiao Ran': A New Ornamental *Lycoris straminea* Cultivar. HortScience. 2023; 58:105-106. https://doi.org/10.21273/HORTSCI16925-22.

3. Zuo Y-W, Quan M-H, Liu G-H, Zhang X, Long N-N, You S-Q, et al. Multi-Omics Analysis Reveals Molecular Responses of Alkaloid Content Variations in Lycoris aurea Across Different Locations. Plant, Cell & Environment. 2024:1-12. https://doi.org/10.1111/pce.15187.

4. Salachna P, Piechocki R. Comparison of nutrient content in bulbs of Japanese red spider lily (*Lycoris radiata*) and golden spider lily (*Lycoris aurea*), ornamental and medicinal plants. World News of Natural Sciences. 2019:75-82.

5. Huang N, Cui L, Wang J, Cai Y, Li X, Xie X, et al. 'Jia Qi Ru Meng': A New Ornamental *Lycoris radiata* × *Lycoris aurea* Cultivar. HortScience. 2024; 59(7):1037-1039. https://doi.org/10.21273/hortsci17943-24.

6. Yamaji F, Ohsawa TA. Breaking-bud pollination: a new pollination process in partially opened flowers by small bees. Journal of Plant Research. 2015; 128(5):803-811. https://doi.org/10.1007/s10265-015-0741-8.

7. Shi T, Yue Y, Shi M, Chen M, Yang X, Wang L. Exploration of Floral Volatile Organic Compounds in Six Typical Lycoris taxa by GC-MS. Plants (Basel). 2019; 8(10). https://doi.org/10.3390/plants8100422.

8. Duke JA, Ayensu ES. Medicinal plants of China. Algonac, Mich.: Reference Publications; 1985.

9. Sun B, Wang P, Wang R, Li Y, Xu S. Molecular Cloning and Characterization of a meta/para-O-Methyltransferase from *Lycoris aurea*. Int J Mol Sci. 2018; 19(7). https://doi.org/10.3390/ijms19071911.

10. Yang Y, Huang S-X, Zhao Y-M, Zhao Q-S, Sun H-D. Alkaloids from the Bulbs of *Lycoris aurea*. Helvetica Chimica Acta. 2005; 88(9):2550-2553. https://doi.org/10.1002/hlca.200590193.

11. Wu Z. Flora Reipublicae Popularis Sinicae, 1 edn: Science Press; 1998.

12. Quan M, Liang J. The influences of four types of soil on the growth, physiological and biochemical characteristics of *Lycoris aurea* (L' Her.) Herb. Scientific Reports. 2017; 7:43284. https://doi.org/10.1038/srep43284.

13. Song J-H, Zhang L, Song Y. Alkaloids from *Lycoris aurea* and their cytotoxicities against the head and neck squamous cell carcinoma. Fitoterapia. 2014; 95:121-126. https://doi.org/10.1016/j.fitote.2014.03.006.

14. Lamoral-Theys D, Andolfi A, Van Goietsenoven G, Cimmino A, Le Calvé B, Wauthoz N, et al. Lycorine, the main phenanthridine Amaryllidaceae alkaloid, exhibits significant antitumor activity in cancer cells that display resistance to proapoptotic stimuli: an investigation of structure-activity relationship and mechanistic insight. Journal of Medicinal Chemistry. 2009; 52(20):6244-6256. https://doi.org/10.1021/jm901031h.

15. Lamoral-Theys D, Decaestecker C, Mathieu V, Dubois J, Kornienko A, Kiss R, et al. Lycorine and its derivatives for anticancer drug design. Mini-Reviews in Medicinal Chemistry. 2010; 10(1):41-50. https://doi.org/10.2174/138955710791112604.

16. Jingting K, Chao J. The mechanism of galanthamine regulating IL-1β/IL-1RA ratio to ameliorate inflammatory microenvironment. Journal of Chinese Pharmaceutical Sciences. 2022; 31(10):773. https://doi.org/10.5246/jcps.2022.10.067.

17. Lima JA, Hamerski L. Alkaloids as Potential Multi-Target Drugs to Treat Alzheimer's Disease. In: Studies in Natural Products Chemistry. Edited by Atta ur R, vol. 61: Elsevier; 2019: 301-334. https://doi.org/10.1016/B978-0-444-64183-0.00008-7.

18. Vezenkov LT, Tsekova DS, Kostadinova I, Mihaylova R, Vassilev NG, Danchev ND. Synthesis of New Galanthamine-Peptide Derivatives Designed for Prevention and Treatment of Alzheimer's Disease. Curr Alzheimer Res. 2019; 16(3):183-192. https://doi.org/10.2174/1567205016666190228123923.

19. Jiang Y, Xia N, Li X, Shen W, Liang L, Wang C, et al. Molecular cloning and characterization of a phenylalanine ammonia-lyase gene (LrPAL) from *Lycoris radiata*. Molecular Biology Reports. 2011; 38(3):1935-1940. https://doi.org/10.1007/s11033-010-0314-9.

20. Liu J-S, Li Y-K, Li J, Li Y, Liu Z-T, Zhou Z-X, et al. Ascorbate peroxidase catalyzes synthesis of protocatechualdehyde from p-hydroxybenzaldehyde in *Lycoris aurea*. Gene. 2024; 927:148697. https://doi.org/10.1016/j.gene.2024.148697.

21. Liang J, Quan M, She C, He A, Xiang X, Cao F. Effects of drought stress on growth, photosynthesis and alkaloid accumulation of *Lycoris aurea*. Pakistan Journal of Botany. 2020; 52(4). http://dx.doi.org/10.30848/PJB2020-4(15).

22. Xu S, Jiang M, Fu J, Liang L, Xia B, Wang R. Physiological and Antioxidant Parameters in Two *Lycoris* Species as Influenced by Water Deficit Stress. HortScience. 2015; 50(11):1702-1708. https://doi.org/10.21273/hortsci.50.11.1702.

23. Zaffar S, Khan I, Zaffar A, Kamili A. Altitudinal Variation in Alkaloid Composition of *Hyoscyamus niger*: A study with reference to Kashmir region of Himalayas. International Journal of Advances in Scientific Research. 2018; 4:13. https://doi.org/10.7439/ijasr.v4i3.4728.

24. Salmore AK, Hunter MD. Elevational Trends in Defense Chemistry, Vegetation, and Reproduction in Sanguinaria canadensis. Journal of Chemical Ecology. 2001; 27(9):1713-1727. https://doi.org/10.1023/A:1010411122739.

25. Zobayed SM, Afreen F, Kozai T. Temperature stress can alter the photosynthetic efficiency and secondary metabolite concentrations in St. John's wort. Plant Physiol Biochem. 2005; 43(10-11):977-984. https://doi.org/10.1016/j.plaphy.2005.07.013.

26. Selmar D, Kleinwächter M. Influencing the product quality by deliberately applying drought stress during the cultivation of medicinal plants. Industrial Crops and Products. 2013; 42:558-566. https://doi.org/10.1016/j.indcrop.2012.06.020.

27. Li Y, Kong D, Fu Y, Sussman MR, Wu H. The effect of developmental and environmental factors on secondary metabolites in medicinal plants. Plant Physiology and Biochemistry. 2020; 148:80-89. https://doi.org/10.1016/j.plaphy.2020.01.006.

28. Popescu R, Kopp B. The genus Rhododendron: An ethnopharmacological and toxicological review. Journal of ethnopharmacology. 2013. https://doi.org/10.1016/j.jep.2013.02.022.

29. Sun Y, Guo R, Geng Y, Shang H, Guo X, Wu Y, et al. Longitudinal Distribution Map of the Active Components and Endophytic Fungi in *Angelica sinensis* (Oliv.) Diels Root and Their Potential Correlations. Metabolites. 2024; 14(1). https://doi.org/10.3390/metabo14010048.

30. Sampaio BL, Edrada-Ebel R, Da Costa FB. Effect of the environment on the secondary metabolic profile of *Tithonia diversifolia*: a model for environmental metabolomics of plants. Scientific Reports. 2016; 6(1):29265. https://doi.org/10.1038/srep29265.

31. Demasi S, Caser M, Lonati M, Cioni PL, Pistelli L, Najar B, et al. Latitude and Altitude Influence Secondary Metabolite Production in Peripheral Alpine Populations of the Mediterranean Species *Lavandula angustifolia* Mill. Frontiers in Plant Science. 2018; 9:983. https://doi.org/10.3389/fpls.2018.00983.

32. Chen S, Liang Z, Webster R, Zhang G, Zhou Y, Teng H, et al. A high-resolution map of soil pH in China made by hybrid modelling of sparse soil data and environmental covariates and its implications for pollution. Science of The Total Environment. 2019; 655:273-283. https://doi.org/10.1016/j.scitotenv.2018.11.230.

33. Dinh QT, Cui Z, Huang J, Tran TAT, Wang D, Yang W, et al. Selenium distribution in the Chinese environment and its relationship with human health: A review. Environment International. 2018; 112:294-309. https://doi.org/10.1016/j.envint.2017.12.035.

34. Chen W, Gong L, Guo Z, Wang W, Zhang H, Liu X, et al. A Novel Integrated Method for Large-Scale Detection, Identification, and Quantification of Widely Targeted Metabolites: Application in the Study of Rice Metabolomics. Molecular Plant. 2013; 6(6):1769-1780. https://doi.org/10.1093/mp/sst080.

35. Fraga CG, Clowers BH, Moore RJ, Zink EM. Signature-Discovery Approach for Sample Matching of a Nerve-Agent Precursor Using Liquid Chromatography−Mass Spectrometry, XCMS, and Chemometrics. Analytical Chemistry. 2010; 82(10):4165-4173. https://doi.org/10.1021/ac1003568.

36. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018; 34(17):i884-i890. https://doi.org/10.1093/bioinformatics/bty560.

37. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology. 2011; 29(7):644-652. https://doi.org/10.1038/nbt.1883.

38. Davidson NM, Oshlack A. Corset: enabling differential gene expression analysis for de novoassembled transcriptomes. Genome Biology. 2014; 15(7):410. https://doi.org/10.1186/s13059-014-0410-6.

39. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nature Methods. 2015; 12(1):59-60. https://doi.org/10.1038/nmeth.3176.

40. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011; 12(1):323. https://doi.org/10.1186/1471-2105-12-323.

41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014; 15(12):550. https://doi.org/10.1186/s13059-014-0550-8.

42. Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. PLoS One. 2016; 11(6):e0157022. https://doi.org/10.1371/journal.pone.0157022.

43. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, et al. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Molecular Plant. 2016; 9(12):1667-1670. https://doi.org/10.1016/j.molp.2016.09.014.

44. Thiel T, Michalek W, Varshney RK, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet. 2003; 106(3):411-422. https://doi.org/10.1007/s00122-002-1031-0.

45. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9(1):559. https://doi.org/10.1186/1471-2105-9-559.

46. González I, Déjean S, Martin PGP, Baccini A. CCA: An R Package to Extend Canonical Correlation Analysis. Journal of Statistical Software. 2008; 23:1-14. https://doi.org/10.18637/jss.v023.i12.

47. Jozefczuk S, Klie S, Catchpole G, Szymanski J, Cuadros-Inostroza A, Steinhauser D, et al. Metabolomic and transcriptomic stress response of *Escherichia coli*. Molecular Systems Biology. 2010; 6:364. https://doi.org/10.1038/msb.2010.18.

48. Guo G, Wang Y, Zhang B, Yu H, Li L, Cao G, et al. Comparative transcriptomic and metabolomic analysis reveals mechanisms of selenium-regulated anthocyanin synthesis in waxy maize (*Zea mays* L.). Frontiers in Plant Science. 2024; Volume 15 - 2024. https://doi.org/10.3389/fpls.2024.1466756.

49. Zhang S, Zhang L, Zou H, Qiu L, Zheng Y, Yang D, et al. Effects of Light on Secondary Metabolite Biosynthesis in Medicinal Plants. Frontiers in Plant Science. 2021; Volume 12 - 2021. https://doi.org/10.3389/fpls.2021.781236.

50. Hussain R, Kim H, Khurshid M, Akhtar M, Linthorst H. Overexpression of AtWRKY50 is correlated with enhanced production of sinapic derivatives in Arabidopsis. Metabolomics. 2018; 14. https://doi.org/10.1007/s11306-018-1317-0.

51. Bi B, Tang J, Han S, Guo J, Miao Y. Sinapic acid or its derivatives interfere with abscisic acid homeostasis during *Arabidopsis thaliana* seed germination. BMC Plant Biology. 2017; 17(1):99. https://doi.org/10.1186/s12870-017-1048-9.

52. Mughal A, Jabeen N, Ashraf K, Sultan K, Farhan M, Hussain MI, et al. Exploring the role of caffeic acid in mitigating abiotic stresses in plants: A review. Plant Stress. 2024; 12:100487. https://doi.org/10.1016/j.stress.2024.100487.

53. Sameiyan E, Hayes AW, Karimi G. The effect of medicinal plants on multiple drug resistance through autophagy: A review of in vitro studies. European Journal of Pharmacology. 2019; 852:244-253. https://doi.org/10.1016/j.ejphar.2019.04.001.

54. Bender O, Atalay A. Polyphenol chlorogenic acid, antioxidant profile, and breast cancer. In: Cancer (Second Edition). Edited by Preedy VR, Patel VB. San Diego: Academic Press; 2021: 311-321. https://doi.org/10.1016/B978-0-12-819547-5.00028-6.

55. Ahammed GJ, Wu Y, Wang Y, Guo T, Shamsy R, Li X. Epigallocatechin-3-Gallate (EGCG): A unique secondary metabolite with diverse roles in plant-environment interaction. Environmental and Experimental Botany. 2023; 209:105299. https://doi.org/10.1016/j.envexpbot.2023.105299.

56. Tu Y, Rochfort S, Liu Z, Ran Y, Griffith M, Badenhorst P, et al. Functional analyses of caffeic acid O-Methyltransferase and Cinnamoyl-CoA-reductase genes from perennial ryegrass (*Lolium perenne*). Plant Cell. 2010; 22(10):3357-3373. https://doi.org/10.1105/tpc.109.072827.

57. Yoder SC, Lancaster SM, Hullar MAJ, Lampe JW. Gut Microbial Metabolism of Plant Lignans: Influence on Human Health. In: Diet-Microbe Interactions in the Gut. Edited by Tuohy K, Del Rio D. San Diego: Academic Press; 2015: 103-117. https://doi.org/10.1016/B978-0-12-407825-3.00007-1.

58. Zhao Y, Hu J, Zhou Z, Li L, Zhang X, He Y, et al. Biofortified Rice Provides Rich Sakuranetin in Endosperm. Rice. 2024; 17(1):19. https://doi.org/10.1186/s12284-024-00697-w.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.docx
- SupplementaryTables.xlsx